Answers to Exam on Topics from Chapters 2 through 13

The model equations are reproduced below.

$$(1) \quad Y_1 = a_1 + a_2 X_1 + a_3 X_2 + a_4 X_3 + u$$
$$(2) \quad Y_2 = b_1 + b_2 Y_3 + b_3 X_2 + v$$
$$(3) \quad Y_3 = Y_2 - Y_1$$

1. The first equation is already in reduced form. Substitute for Ys from the first and second equations into the third. We have,

$$\begin{aligned} Y_3 &= b_1 + b_2 Y_3 + b_3 X_2 + v - (a_1 + a_2 X_1 + a_3 X_2 + a_4 X_3 + u) \\ &= (b_1 - a_1) + b_2 Y_3 - a_2 X_1 + (b_3 - a_3) X_2 - a_4 X_3 + (v - u) \end{aligned}$$

Solving for Y_3, we obtain the reduced form for Y_3 as follows.

$$(4) \quad Y_3 = \frac{(b_1 - a_1)}{(1 - b_2)} - \frac{a_2}{(1 - b_2)} X_1 + \frac{(b_3 - a_3)}{(1 - b_2)} X_2 - \frac{a_4}{(1 - b_2)} X_3 + \frac{(v - u)}{(1 - b_2)}$$

$$= p_1 + p_2 X_1 + p_3 X_2 + p_4 X_3 + e_1$$

The reduced form for Y_2 is obtained by substituting for Y_3 from the above equation into the second equation.

$$\begin{aligned} (5) \quad Y_2 &= b_1 + b_2(p_1 + p_2 X_1 + p_3 X_2 + p_4 X_3 + e_1) + b_3 X_2 + v \\ &= (b_1 + b_2 p_1) + b_2 p_2 X_1 + (b_2 p_3 + b_3) X_2 + b_2 p_4 X_3 + e_2 \\ &= m_1 + m_2 X_1 + m_3 X_2 + m_4 X_3 + e_2 \end{aligned}$$

2. Because the first equation does not involve any endogenous variables that are correlated with the error terms, we can apply OLS estimation and obtain estimators that are BLUE.

3. The second equation has Y_3 on the right hand side, which is correlated with v (as seen from the reduced form equation for Y_3). Therefore, OLS estimators are biased and inconsistent.

4. First estimate the reduced from for Y_3 by regressing it against a constant, X_1, X_2, and X_3 and save the predicted value as $\hat{Y}_3 = \hat{p}_1 + \hat{p}_2 X_1 + \hat{p}_3 X_2 + \hat{p}_4 X_3$. Next regress Y_2 against a constant, $\hat{Y}_3$, and X_2.

5. The first equation is already in reduced form and hence it is uniquely estimable. Since $m_2 = b_2 p_2$, we can estimate b_2 as $\hat{m}_3/\hat{p}_2$. However, this is not unique because we also have $\tilde{b}_2 = \hat{m}_4/p_4$. Because b_1 and b_3 depend on b_2, their estimates too are not unique. This is because the second equation is over-identified and hence TSLS is the appropriate estimation procedure for that equation.